The Interaction of Coarticulation and Prosody in Sound Change

Mary E. Beckman, Kenneth De Jong, Sun-Ah Jun and Sook-Hyang Lee
Language and Speech 1992 35: 45
DOI: 10.1177/002383099203500205

THE INTERACTION OF COARTICULATION AND PROSODY IN SOUND CHANGE*

MARY E. BECKMAN,
Ohio State University
KENNETH DE JONG
University of California, Los Angeles
SUN-AH JUN and SOOK-HYANG LEE
Ohio State University

Ohala (1974, 1981a) has proposed that sound changes can originate in hearers' misinterpretations of synchronic phonetic patterns. This paper applies this idea to sound changes that are conditioned by the prosodic environment, such as the voicing of voiceless fricatives in unstressed syllables in Proto-Germanic. Browman and Goldstein's (1989, 1990) "gestural score" suggests a representation of synchronic patterns in which extreme overlap between gestures of neighboring phoneme segments in casual speech can produce the appearance of a feature change or a segment deletion. Many of the sound changes that are conditioned by prosodic environment can be viewed as a diachronic reinterpretation of just such synchronic fast-speech processes. For example, vowel reduction in unstressed syllables can be viewed as a reinterpretation of undershoot that occurs when the vowel is overlapped to a great extent by the oral gestures for neighboring consonants. Phonetic data are reviewed that support analogous accounts of stop spirantization, voiceless obstruent voicing, and even the insertion of an intrusive stop in clusters such as /ns/ in some prosodic environments.

Key words: sound change, coarticulation, stress, gestural overlap

INTRODUCTION

Among the many lasting contributions that John Ohala has made to our field, one of the most important is his idea that sound changes often originate in listeners' reinterpretations of ordinary synchronic sound patterns, patterns such as the coarticulatory influences of neighboring segments upon each other (Ohala, 1974, 1981a). This idea puts phonetic explanations for sound change squarely in the realm of experimental verification. For example, by examining the effects of a following nasal on a vowel's

* This work was supported in part by the National Science Foundation under grant IRI-8858109. We thank Neal Jacobs for the example of intrusive stops from Yiddish, and Brian Joseph for extensive help and advice with many of the other examples of prosodically conditioned sound changes. We thank John Ohala for being such an inspiring teacher of our first phonetics course in the fall of 1976, or an inspiring role model for our teacher in 1985 and 1986.

Send correspondence to Mary Beckman at Ohio State University, Linguistics Department, 222 Oxley Hall, 1712 Neil Ave., Columbus, OH 43210-1298, USA.

formants and by testing how listeners perceive these effects when the vowel is spliced out of its original consonantal context, we can surmise that the neutralization of high and mid nasalized vowels in French and other languages arose when listeners misinterpreted the acoustic effects of coarticulatory nasalization at a time when syllable-final nasal consonants were being weakened on the way to being elided altogether (Ohala, 1974; Beddor, Krakow, and Goldstein, 1988; Krakow, Beddor, Goldstein, and Fowler, 1988). Such explanations for sound change, however, are only as good as the phonetic model of the synchronic sound pattern which the listener (mis)interprets.

In this paper, we will talk about one type of phonetic representation that we think is useful for understanding sound changes that are conditioned by the prosodic context. The appendix lists several such examples, which at first glance may seem to have little else in common. How can a single phonetic explanation account for the prosodically conditioned exceptions to the voicing of the medial fricative in Proto-Germanic and to the insertion of a stop between a lateral and a sibilant in Yiddish? We will argue that, if the right phonetic model is used to represent coarticulatory interactions between the neighboring segments in the historically attested or reconstructed sound sequences, many of these exceptions can be modeled in terms of the effect of the prosodic environment on some common aspect of their production.

Coarticulation in the gestural score

The representation that we have in mind is one in which phoneme sequences are not modeled as strictly sequential bundles of simultaneous feature targets. Rather, each phoneme segment involves the orchestration of several more or less simultaneous gestures, each of which can overlap more or less with the gestures of neighboring segments in a complex gestural score. When the overlapped gestures of neighboring segments involve different articulators, the resulting acoustic patterns will reflect the aerodynamic consequences of a vocal tract shaped by all of the simultaneous constrictions. When they involve some articulators in common, on the other hand, the observed articulatory constrictions themselves will reflect a compromise or "blending" of the demands of the different gestures. Thus, for example, in a /su/ or /si/ sequence, the vowel's labial gesture overlaps with the consonant's apico-alveolar constriction and glottal opening gestures, so that the frication noise is spectrally colored by the concurrent rounding or spreading. In a /kV/ sequence, the consonant's dorsovelar constriction blends with the overlapping dorsal gesture for the vowel to produce the widely differing places of articulation and extremely variable release spectra observed for /k/ in different vocalic environments.

This idea of coarticulation as overlap and blending of gestures was formulated quantitatively in Öhman's pioneering work in the 1960s (Öhman, 1966, 1967). More recently, it has been developed extensively within a task-dynamic framework by authors such as Saltzman and Munhall (1989), Browman and Goldstein (1990), Boyce, Krakow, Bell-Berti, and Gelfer (1990), and Munhall and Löfqvist (1992). Browman and Goldstein (1989) have applied this model to propose a unified explanation for seemingly disparate processes of segment weakening, assimilation, and deletion in casual speech. They give an example of a pair of utterances of the phrase *perfect memory*, one in which a pause intervenes between the two words and another more fluent rendition in which the final /t/ of *perfect* seems to have been deleted. However, X-ray microbeam pellet traces show that both renditions have a clear apico-alveolar gesture for the /t/. The only difference is that in the fluent rendition the end of the gesture is hidden aerodynamically and acoustically by the overlapping labial gesture for the following /m/, whereas the lack of overlap in the rendition with a pause allows the pent-up air behind the alveolar closure to be audibly released.

An analogous example of an apparent assimilation is discussed in De Jong (1991). He examined X-ray microbeam pellet tracings during the /utðə/ sequence in repetitions of the utterance *(I said) Put the toe on the table,* with two different stress patterns: In one case the stress was on the first word, *put,* and in the other the stress was on *toe.* In utterances with nuclear stress on the first word, there was a brief pause after the accented *put,* during which the tongue moved forward from the alveolar [t] to the dental [ð]. In utterances with stress on *toe,* by contrast, there was no pause after the *put,* and here the trajectory for the tongue blade pellet suggested a single gesture for the place of the two constrictions in the consonant cluster, at a site considerably anterior to that in the [t] but not so front as in the [ð] of the other utterances. A gestural score representation allows us to model this utterance as a case of extreme gestural overlap, just like the apparent deletion of the /t/ in Browman and Goldstein's example. If we were to model the utterance in terms of the discrete sequential representation of the International Phonetic Alphabet (IPA), on the other hand, we would have no recourse other than to transcribe an allophonic feature change: /tð/ → [t̪ ð]. Browman and Goldstein (1989) propose that many casual speech rules are essentially like these two cases of apparent assimilation and deletion; the linguist-observer interprets the various results of closely overlapped and blended gestures in terms of segment deletions or qualitative feature changes in order to make a symbolic transcription.

Here, we will sketch a proposal for a similar explanation of prosodically conditioned sound change, with the ordinary listener's everyday phonological interpretations replacing those of the transcribing phonetician, as in Ohala's picture of common sound changes. We predict that any prosodic effect that increases the gestural overlap, or that decreases the acoustic salience of an overlapped gesture, would increase the likelihood of a listener reinterpreting the coarticulation as an intentional feature of the affected phoneme segment. For example, an unstressed vowel might be very short, so that a greater proportion of its dorsovelar gesture overlaps with the preceding consonant. A listener might misinterpret the resulting coarticulation as an intentional vagueness about the vowel's quality – that is, an underlying full vowel might be replaced with /ə/, as in English /dʒɪmnəst/ for *gymnast.* Or, in the case of more extreme overlap, the listener might misinterpret the short vowel as nothing more than the consonant's release – that is, a syllable might be lost, as in English /tʃɑklət/ for *chocolate.* Thus, reduction and deletion of an unstressed vowel can be represented plausibly as two steps on a continuum of increasing overlap between the vowel and its neighboring consonants. We will illustrate our general proposal by first reviewing some evidence for this account of vowel reduction and deletion, and then suggesting analogous gestural-score representations for some of the other prosodically conditioned sound changes listed in the appendix.

REDUCTION AS TRUNCATION OF GESTURES

Vowel reduction

Edwards, Beckman, and Fletcher (1991) present some synchronic data that are compatible with the model of vowel reduction just proposed. In this study of jaw kinematics in prosodically various /pɑp/ sequences, they found that the jaw moved a shorter distance when lowering for shorter unaccented vowel tokens than for longer accented ones. The other kinematic characteristics of the movements, however, did not suggest a higher target for the unaccented vowel, leading Edwards *et al.* to interpret the smaller displacement values in the unaccented syllables as a "truncation" effect of the sort that Bullock and Grossberg (1988) have observed for limb movements. That is, the jaw-raising movement into the labial closure for the following /p/ cuts off the oral gesture for the vowel before the jaw reaches its intended low position. From the point of view of the speaker, this effect is an artifact of properties of the syllable other than the intended vowel quality. From the point of view of the listener, however, there might be little difference between a truncated low vowel and an intentionally higher vowel. A listener who misinterprets the truncation as an intended feature of the vowel quality might generalize a rule replacing /ɑ/ in unaccented syllables with /ə/.

This account of historic vowel reductions calls to mind Lindblom's (1963) idea of target undershoot as the source of synchronic reduction of short vowels. Lindblom assumed that a sequence of phonemes, such as a CVC syllable, is produced as a sequence of temporally invariant movements toward the successive segmental targets. If the movement toward the final consonant begins before the vowel target is reached, the formants during the vowel will be more similar to their locus values for the neighboring consonants, reducing distinctions among the vowels in a given consonantal context. In the gestural score model, the situation is more complicated, because gestures are not temporally invariant movements, and because gestures on different tiers may be co-ordinated differently. However, overlap among gestures can produce undershoot on one or more articulatory tiers. In the CVC sequence, the tongue-body gesture for the vowel shares at least one articulator with the oral gestures of neighboring consonants; because of the physical coupling between the lower lip and jaw and between the tongue tip and tongue body, the vowel gesture shares the jaw with labial consonant gestures and the tongue body itself with coronal and velar consonant gestures. In the conditions that lead to a shorter observed vowel duration, these gestures would overlap more and require extensive blending when the common articulators are marshalled to produce the gesture. At shorter and shorter durations, more and more of the vowel's gesture would be blended with the following consonant gesture, leading to more and more truncation of the vowel target. Although later studies such as Harris (1978) highlight the residue, Lindblom's simple undershoot model did account for a great deal of the variation in vowel formant values just from the observed duration values and the fixed frequency values chosen for the different underlying vowel targets and context consonant loci. Thus, although we would now say that other facts about gestural organization and timing are essential for a full understanding, truncation of closely phased gestures seems to account for a great deal of vowel reduction synchronically.

Consonant lenitions

Historical linguists have traditionally suggested a phonetic affinity between such vowel reduction and various sound changes that affect consonants by calling them all "lenitions" or "weakenings". For example, Bloomfield (1922) uses the term in describing the flapping of alveolar stops in American English, the spirantization of Latin voiced stops in Spanish, and Verner's law (i.e., the voicing of fricatives in unaccented syllables in ProtoGermanic). The idea of an affinity among these various sound changes probably stems from the observation that they are often limited to occurring in weak prosodic positions. Vowel reductions and deletions are typical of unstressed syllables. Verner's law, similarly, is interpreted as evidence that the ProtoGermanic reflex of the ProtoIndo-European accent involved a contrast in syllable strength – that is, it already had become the stress-accent type that is found in Modern German, Danish, English, and so on (Verner, 1875/1967, p. 160). We might ask, therefore, whether this conventional wisdom has any physical basis. Is there an underlying phonetic affinity? In particular, is there anything about stop spirantization and obstruent voicing that suggests that both processes can be modeled as being essentially similar to vowel reduction?

Spirantization of stops

The spirantization of a stop consonant in prosodically weak positions looks very much like the converse of vowel reduction – namely, a truncation and apparent reduction of the oral gesture for the consonant. Again, we observe that the oral closure gesture for the consonant shares at least one articulator with the tongue-body gesture for the neighboring vowels. In prosodic contexts where segment durations are generally shorter, the shared components of the consonant's oral gesture might be truncated by the overlapped vowel gestures. If the undershoot results in a less perfect seal, a seal through which air can leak, then either turbulence may ensue during the closure as air pressure builds up behind the leak, or enough air pressure may be vented noiselessly through the leak to prevent a strong stop burst on release. In the first case, the listener might misinterpret the turbulence as intentional frication. In the second, the listener might misinterpret the lack of a strong burst as an intentional feature of a weak non-sibilant fricative or even of an approximant.

In Lindblom's (1963) data that we cited above as evidence for vowel truncation, there was no evidence of this converse truncation of the consonant in shorter syllables. In Edwards *et al.*'s (1991) data likewise, there was no consistent pattern of lower jaw values for /p/ in unaccented syllables. This lack of evidence for consonant reduction in two sets of data that show vowel reduction suggests a greater resistance of stop consonants to undershoot, which in turn may reflect a stability inherent to the contact between opposed articulators. For example, bilabial closing gestures can accommodate a great deal of variation in the force with which the lower lip is raised or the upper lip lowered, because each articulator is braced against the other opposing articulator. The lower lip might be more and more forcibly pressed against the upper lip, but it cannot physically move through it. Stone, Faber, Raphael, and Shawker (1992) suggest that, in many lingual consonants, muscle forces are marshalled to allow similar bracings of the tongue surface against the opposing lateral surfaces of the palate. While such stability due to bracing might figure in the articulation of some vowels (*cf.* Fujimura and Kakita, 1979; Perkell and Nelson, 1982, 1985), consonants as a class, and stop consonants in particular, would seem to be naturally more resistant to truncation of their oral gestures, suggesting that audible undershoot of stop closure gestures should be somewhat less frequent than the ubiquitous undershoot of tongue-body configurations for vowels.

On the other hand, there are other more anecdotal data which do show the synchronic reduction of stops to fricatives or continuants. For example, Fant and Kruckenberg (1989, p. 2) note, in a database of short read texts in Swedish, "frequent incomplete oral closure of voiced stops and nasals which become reduced to voiced continuants and nasalized vowels". We similarly have observed noticeable frication in productions of intervocalic /b/ and /d/ in our own data from Tokyo Japanese, particularly data from Shitamachi speakers.

As a diachronic process, the spirantization (or vocalization) of stops is fairly common. Indo-European alone readily yields examples throughout its history. As part of the First Germanic Sound Shift (a.k.a. Grimm's Law), ProtoIndoEuropean (PIE) *p, *t, *k became ProtoGermanic *f, *θ, *x in all contexts except clusters after a preceding voiceless consonant. ProtoGermanic *p, *t, *k (from PIE *b, *d, *g by Grimm's Law) then became, by the Second Germanic Sound Shift, affricates /pf, ts, kx/ in root-initial position, and fricatives /f, s, x/ elsewhere. The Ancient Greek voiceless aspirated stops /ph, th, kh/ were replaced by fricatives /f, θ, x/, starting in the late classical period. At about the same time (but at an independent rate), the Classical Greek voiced stops /b, d, g/ were replaced by /β, ð, γ/. Voiced stops from Latin regularly have gone to voiced fricatives word-medially between vowels in Iberian Spanish and in many American dialects of Spanish. Voiced and voiceless stops in Old Irish became fricatives intervocalically, both word-medially and at phrase-medial word boundaries after prosodically weak forms such as pronouns and auxiliary verbs. The change in the latter prosodic environment led to the highly irregular morphological alternations of Modern Irish when the fricatives at word boundaries were reanalyzed as inflectional marks of person, tense, and aspect.

As these examples show, however, spirantization is not necessarily limited to weak prosodic contexts. While our impression is that the position-conditioned changes in Romance and Celtic are more typical, many of the cases just cited are not conditioned in this way. Perhaps there is an alternative phonetic source of spirantization that does not stem from truncation. The two sound changes affecting /ph, th, kh/ in Greek and *p, *t, *k in Proto-Germanic are suggestive. We note first that the exception of the cluster environment for Grimm's Law supports the hypothesis that the voiceless stops were generally aspirated in Proto-Germanic at the time that they underwent spirantization, so that this change probably is the same as the Greek one (see, e.g., Prokosch, 1939, p. 59). Might a listener misinterpret the turbulence of the aspirated release as frication, even in the absence of truncation of the oral gesture for the stop?

If there are two such potentially independent phonetic sources for the diachronic spirantization of stops, we can make the following predictions about them. Misinterpretation of truncated closures (i.e., "true lenition") should be more typical of voiced stops. We predict this because voiced stops typically have shorter closure durations (see, e.g., the survey in Lehiste, 1970). Also, kinematic data generally show that closing movements into voiced stops have smaller velocities for the same duration and have lower jaw displacements than those for voiceless stops (see, e.g., Summers, 1987), suggesting a less forceful, less braced gestural target. Therefore, spirantization of stops only in prosodically weak position should be more common for voiced than for voiceless stops. By contrast, reanalysis of release turbulence would be most typical of strongly aspirated voiceless stops. Moreover, in strong prosodic positions, where the stop closure is less likely to undergo truncation, affricates could be the end result of the sound change instead of fricatives. This last prediction is in accord with the prosodically conditioned reflexes of the Second Germanic Sound Shift in Old High German.

Voicing of stops and fricatives

Like the spirantization of stops in weak prosodic position, the voicing of intervocalic obstruents can be modeled as a truncation of the consonant gesture by the gestures of neighboring vowels, except that in this case the salient acoustic effect involves the glottal tier rather than tiers in the oral subsystem. Evidence for gestural blending on the glottal tier independent of the other tiers can be found in a study by Munhall and Löfqvist (1992). They tracked glottal width over time in utterances of the sentence *Kiss Ted*, produced at many different speech tempi. In very slow tempo tokens they observed complete glottal opening and closing movements for each of the two medial consonants. That is, there were two distinct peaks in the glottal width trace, one in the middle of the /s/ constriction and the other around the release of the /t/, separated by a clear trough where the glottal width was zero. At normal rate, the closing of the glottis for /s/ overlapped with the opening of the glottis for /t/; the trough diminished or disappeared while preserving two fairly distinct peaks, or a shoulder followed by a peak. At faster and faster rates, the two peaks came closer and closer together, until at the fastest rate, there was only a single smooth abduction and adduction, with the peak located at the oral closure into the stop, as it is in syllable-initial /st/ clusters with unaspirated [t] (as if the second word were *stead* instead of *Ted*).

When we interpret data on the time course of glottal width for a single intervocalic voiceless stop consonant in light of this extensive study of a consonant sequence, we see evidence for similar blending between the abduction gesture for the consonant and the adduction gestures for the neighboring vowels. For example, Dixit (1989) shows glottal opening and closing movements in voiceless aspirated stops in Hindi to be shorter and the peak opening to be greatly reduced in word-medial as compared to word-initial position. Generalizing from this, we might expect a similar shortening and truncation for word-medial unaspirated stops.[1] If airflow conditions were right, this reduced glottal

[1] Dixit's single subject actually showed an even more drastic reduction, producing most of the word-medial /p/ tokens with no evidence of any glottal opening. This contrasts with subjects in other studies, such as that of Benguerel and Bhatia (1980), which show a sequence of glottal opening and closing movements in word-medial voiceless unaspirated stops. In this earlier study, however, it is not possible to compare the time course and peak value with comparable word-initial productions, because the forms were produced in isolation, so that the initial stops showed only a closing gesture from a wide-open glottis for respiration.

opening would allow voicing to continue some of the way into the stop closure, which might then be misinterpreted as an intentional feature of the stop.

If intervocalic voicing is the result of such truncation of the glottal gesture for a voiceless obstruent, we can predict that it would occur in just those environments where a consonant's glottal gesture tends to be truncated in general. Thus, if a language has both unaspirated and aspirated stops, we would expect voicing of the unaspirated stops to occur in just those prosodic contexts where voice onset time is shorter in the aspirated stops. This is true of Korean. Jun (1990) has shown that VOT is shorter for the aspirated stops in just the same prosodic positions where the lenis-stop voicing rule applies – namely, at word boundaries within an accentual phrase, but not across accentual-phrase boundaries. The positional variants of phonemically voiced and voiceless stops in American English also accord with this prediction; /b, d, g/ are more consistently voiced intervocalically in the onset of unstressed syllables, which is where /p, k/ are not aspirated and /t/ is usually replaced by [ɾ]. The same distributional pattern holds for Mandarin Chinese. The voiceless unaspirated stops /p, t, k/ are often voiced at the beginning of the second syllable in words such as [li^{55} ba] /lī pa/ 'fence'. Here the second syllable has no tonal specification and therefore counts as unstressed. (Contrast [ti^{51} pa^{55}] /tì pā/ 'number 8', where the phonological specification of a high level tone on the segmentally identical syllable constitutes stress, and the /p/ does not voice.) The Armenian and Greek sound changes shown in the appendix are probably similar. The forms in which PIE *t after *n went to /d/ seem to be cases where the stop is initial to an unstressed syllable after a stressed syllable.

The ProtoGermanic sound change, on the other hand, is the opposite; the voiceless fricatives remained voiceless in sequences of syllables with a falling stress pattern (where flapping occurs in American English), and became voiced between an unstressed vowel and a stressed vowel. However, here the syllable structure was different. In his classic article describing this sound change, Verner (1875/1967, p. 149) pointed out that alliteration rules in poetry not long after the change show that the fricatives must have been codas for the preceding syllable rather than onsets for the following syllable. The fricatives that did not become voiced thus belonged to stressed syllables. They were probably longer and their glottal gestures overlapped proportionally less with the gestures for the neighboring vowels. Conversely, the fricatives that were voiced might be shorter and their gestures overlapped proportionally more. Moreover, if their gestures on the oral tier were also somewhat truncated, the oral air pressure behind the leakier constriction would be vented to some extent, allowing greater airflow through the glottis to maintain voicing even as the vocal folds are abducting.

Epenthetic Stops

Unlike the sound changes described in the previous section, stop epenthesis is already traditionally described in terms of a change in the pattern of overlap among the component gestures in the sequence. For example, Bloomfield (1922, pp. 383–384) accounts for the insertion of a voiced stop after the nasal in words such as *thunder* (OE *þunor*) and *thimble* (OE *þymel*) as follows:

> These changes involve no additional movement, but merely replace simultaneous movements by successive. To pass from [n] to [r], for instance, the speaker must simultaneously raise his velum and move his tongue from the closure position to the trill position.... If, with a less delicate co-ordination, the velum is raised before the change of tongue-position, there results a moment of unnasalized closure, equivalent to the phoneme [d].

Replacing "the velum is raised" in this account with "the sides of the tongue are raised" yields a similar hypothesis about the insertion of the [d] into the /lr/ sequence in *alder* (OE *alor*). The analogous hypotheses about the synchronic insertion of [t] in nasal-fricative or lateral-fricative sequences in words such as *dance* or *false* are also commonplace; each has been described as a delay in the achievement of the narrow apico-alveolar channel of the fricative relative to the closing off of the velopharyngeal or lateral lingual air passage of the preceding consonant (e.g., Ohala, 1974, 1981b). What is not clear in these accounts, however, is why there should be this delay. In particular, it is not clear why there should be a greater likelihood of such a delay in a tautosyllabic cluster or in a stressed syllable, as in the sound changes in the appendix.

However, this delay is not the only temporal reorganization that can insert an apparent stop closure into such sequences. Ali, Daniloff, and Hammarberg (1979) list at least two additional mechanisms that might create a short silent gap in a nasal-fricative cluster, and the mechanism that seemed best supported by their airflow and pressure data was the opposite of the one traditionally assumed. That is, suppose that the forming of the fricative channel is not late relative to the closing off of the velopharyngeal port, but rather is early. The continued leakage of air out of the nose would retard the buildup of sufficient oral air pressure to start turbulent flow through the fricative constriction. This aerodynamic pattern might occur naturally as the result of overlap and blending between the gestures for the nasal and the fricative, which require different degrees of constriction at the tongue tip and diametrically opposite configurations for velic aperture. When the gestures involving the same articulatory system blend together, there might result a short interval during which the tongue blade has already lowered away from the alveolar ridge to form the narrow channel of the fricative constriction while the velopharyngeal port is still slightly open, allowing just enough leakage to offset the airflow from the opening glottis. In a cluster with a lateral, similarly, the /l/ and the fricative require opposite configurations for the tongue body. In the /l/, the tongue body is raised medially and lowered laterally to channel air out of the sides, whereas in the /s/, it is raised laterally and deeply grooved medially to channel air centrally into the anterior slit (see, e.g., Stone *et al.* 1992). The blending of the two tongue body gestures might allow a slight leakage of air laterally, venting air pressure away from the central anterior opening.

The likelihood of such an aerodynamic result would depend on the degree of overlap between the two gestures. When proportionally less of the nasal or lateral gesture overlaps with the fricative, there would be a lesser likelihood of getting nasal or lateral leakage from the blending. Browman and Goldstein's (1989) observations of gestural timing within and across syllables suggest that the greatest degree of overlap will occur when the two consonants are both in the onset or both in the coda of the same syllable, as in *mince* or *false*.[2] Here there should be the greatest likelihood of the coarticulation being

[2] An example of epenthesis in onset clusters is the insertion of /t/ in Germanic and Slavic for PIE initial *sr-, as in *stream* (*cf.* Sanskrit, *sravati* 'flows').

misinterpreted as an intentional stop. In the heterosyllabic clusters of *dancing* or *falsity*, by contrast, a much smaller proportion of the sonorant consonant would overlap with the following syllable-initial obstruent, reducing the likelihood of an epenthetic stop. In the unstressed syllable in the Yiddish example, similarly, there should be less drastic overlap because here the nasal or lateral consonant is not part of the coda with the obstruent, but instead is functioning as the syllable's nucleus.

CONCLUSION

In this paper, we have reviewed some evidence in support of the idea that when coarticulation is represented in terms of gestural overlap, it accords well with John Ohala's account of sound change as the misinterpretation of common synchronic sound patterns. When viewed this way, the prosodic conditioning of some sound changes seems to fall out from the predicted effects of different prosodic positions on the likelihood and extent of overlap on different articulatory tiers.

REFERENCES

ALI, L.H., DANILOFF, R.G., and HAMMARBERG, R. (1979). Intrusive stops in nasal-fricative clusters: An aerodynamic and acoustic investigation. *Phonetica*, **36**, 85–97.

BEDDOR, P.S., KRAKOW, R.A., and GOLDSTEIN, L.M. (1988). Perceptual constraints and phonological change: A study of nasal vowel height. *Phonology Yearbook*, **3**, 197–217.

BENGUEREL, A.-P., and BHATIA, T.K. (1980). Hindi stop consonants: An acoustic and fiberscopic study. *Phonetica*, **37**, 134–148.

BLOOMFIELD, L. (1922). *Language*. New York: Holt, Rinehart, and Winston.

BOYCE, S.E., KRAKOW, R.A., BELL-BERTI, F., and GELFER, C.E. (1990). Converging sources of evidence for dissecting articulatory movements into core gestures. *Journal of Phonetics*, **18**, 173–188.

BROWMAN, C.P., and GOLDSTEIN, L. (1989). Tiers in Articulatory Phonology, with some implications for casual speech. In J. Kingston and M.E. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech* (pp. 341–376). Cambridge, U.K.: Cambridge University Press.

BROWMAN, C.P., and GOLDSTEIN, L. (1990). Gestural specification using dynamically defined articulatory structures. *Journal of Phonetics*, **18**, 299–320.

BULLOCK, D., and GROSSBERG, S. (1988). The VITE model: A neural command circuit for generating arm and articulator trajectories. In J.A.S. Kelso, A.J. Mandell, and M.S. Shlesinger (eds.), *Dynamic Patterns in Complex Systems* (pp. 305–326). Singapore: World Scientific Publishers.

DE JONG, K. (1991). *The Oral Articulation of English Stress Accent*. Doctoral dissertation, Ohio State University.

DIXIT, P. (1989). Glottal gestures in Hindi plosives. *Journal of Phonetics*, **17**, 213–237.

EDWARDS, J., BECKMAN, M.E., and FLETCHER, J. (1991). The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America*, **89**, 369–382.

FANT, G., and KRUCKENBERG, A. (1989). Preliminaries to the study of Swedish prose reading and reading style. *Speech Transmission Laboratory Quarterly Progress Report* 1989, **2**, 1–83. (Stockholm: Kungliga Teknika Högskolan.)

FUJIMURA, O., and KAKITA, Y. (1979). Remarks on quantitative description of the lingual articulation. In B. Lindblom and S. Öhman (eds.), *Frontiers of Speech Communication Research* (pp. 17–24). London: Academic Press.

HARRIS, K.S. (1978). Vowel duration change and its underlying physiological mechanisms. *Language and Speech*, **21**, 354–361.

HOCK, H.H. (1976). Review of Raimo Antilla, *An Introduction to Historical and Comparative Linguistics. Language*, **52**, 202–220.

JUN, S.-A. (1990). The domain of laryngeal feature effects in Chonnam Korean. *Journal of the Acoustical Society of America*, **87**, S123.

KRAKOW, R.A., BEDDOR, P.S., GOLDSTEIN, L.M., and FOWLER, C.A. (1988). Coarticulatory influences on the perceived height of nasal vowels. *Journal of the Acoustical Society of America*, **83**, 1146–1158.

LEHISTE, I. (1970). *Suprasegmentals.* Cambridge, MA: MIT Press.

LINDBLOM, B. (1963). A spectrographic study of vowel quality. *Journal of the Acoustical Society of America*, **35**, 1773–1781.

MUNHALL, K., and LÖFQVIST, A. (1992). Gestural aggregation in speech: Laryngeal gestures. *Journal of Phonetics*, **20**, 111–126.

OHALA, J.J. (1974). Experimental historical phonology. In J.M. Anderson and C. Jones (eds.), *Historical Linguistics II: Theory and Description in Phonology* (pp. 353–389). Amsterdam: North-Holland.

OHALA, J.J. (1981a). The listener as a source of sound change. In C.S. Masek, R.A. Hendrick, and M. Miller (eds.), *Papers from the Parasession on Language and Behavior* (pp. 178–203). Chicago: Chicago Linguistics Society.

OHALA, J.J. (1981b). Speech timing as a tool in phonology. *Phonetica*, **38**, 204–212.

ÖHMAN, S.E.G. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, **39**, 151–168.

ÖHMAN, S.E.G. (1967). Numerical model of coarticulation. *Journal of the Acoustical Society of America*, **41**, 310–320.

OLSEN, B.A. (1989). A trace of Indo-European accent in Armenian. *Historische Sprachforschung*, **102**, 220–240.

PERKELL, J.S., and NELSON, W.L. (1982). Articulatory targets and speech motor control: A study of vowel production. In S. Grillner, A. Persson, B. Lindblom, and J. Lubker (eds), *Speech Motor Control* (pp. 187–204). New York: Pergamon Press.

PERKELL, J.S., and NELSON, W.L. (1985). Variability in production of the vowels /i/ and /a/. *Journal of the Acoustical Society of America*, **77**, 1889–1895.

PROKOSCH, E. (1939). *A Comparative Germanic Grammar (William Dwight Whitney Linguistics Series).* Baltimore, MD: Linguistic Society of America.

SALTZMAN, E.L., and MUNHALL, K.G. (1989). A dynamical approach to gestural patterning in speech. *Ecological Psychology*, **1**, 333–382.

STONE, M., FABER, A., RAPHAEL, L.J., and SHAWKER, T.H. (1992). Cross-sectional tongue shape and lingua-palatal contact patterns in [s], [ʃ], and [l]. *Journal of Phonetics*, **20**, 253–270.

SUMMERS, W.V. (1987). Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *Journal of the Acoustical Society of America*, **82**, 847–863.

VERNER, K. (1875/1967). An exception to the first sound shift. English translation in W.P. Lehmann (ed.), *A Reader in 19th Century Historical Indo-European Linguistics* (pp. 132–163). Bloomington, IN: Indiana University Press.

APPENDIX

Prosodically conditioned sound changes

Vowel reductions and deletions

1. At various stages in the history of English, there have been many reductions and/or deletions of unstressed vowels, resulting in the whole-scale deletion of final /ə/, of the vowel in plural endings, and so on:

/sɪŋ/	*sing*	<	OE *singe*
/stonz/	*stones*	<	OE *stanas*
/mʌŋks/	*monks*	<	OE *munecas*

The conditioning by stress is particularly clear in doublets such as /hauswaɪf/ *housewife* versus /hʌzəf/ (<*housewife*) = 'sewing kit', /ɑn bord/ *on board* versus /əbord/ *aboard* (< *on+board*). Synchronically, there are similar changes in some speakers' productions of reduced vowels in syllables originally having only secondary stress and of no vowel in medial completely unstressed syllables, as in many people's pronunciations of the following and other words:

/ɹəkun/	<	/ɹækun/	'raccoon'	/kjugə/	<	/kəjugə/	'Cayuga'
/dʒɪmnəst/	<	/dʒɪmnæst/	'gymnast'	/tʃɑklət/	<	/tʃɑkələt/	'chocolate'

2. French underwent a drastic deletion of unstressed syllables in its evolution from Vulgar Latin. Compare Modern French *manche* 'sleeve' < Vulgar Latin *mánica* with Modern French *tenèbres* 'darkness' < Vulgar Latin *tenébras.*

3. Unstressed vowels were deleted in Old Irish in the medial syllable immediately following the primary stress. This change is traditionally analyzed as conditioned by distance from the beginning of the word rather than by stress. Hock's (1976) reanalysis in terms of position relative to the stress is based on examples such as the following, where stress was not initial and no vowel was deleted or some other vowel was deleted instead:

asbérat	<	*essbérodd	'they speak'
arna-ébret	<	-*éssbérodd	'they do not speak'

Obstruent lenitions – spirantization (and vocalization) of stops

1. Second Germanic Sound Shift. In Old High German, the Proto-Germanic voiceless stops *p *t *k (which earlier developed from Proto-IndoEuropean *b *d *g by Grimm's Law) became affricates /pf, ts, kx/ in root-initial position and fricatives /f, s, x/ elsewhere, as in:

initial /pf, ts/	non-initial /f, s/ < *p *t
/pfat/ *Pfad* 'path'	/ʃlafn/ *schlafen* 'sleep'
/tsait/ *Zeit* 'time'	/aws/ *aus* 'out'

Note: Since stress was considerably root-initial at the time of the shift, the conditioning could also be described as affricate at the beginning of stressed syllables, fricatives elsewhere.

2. In many dialects of Spanish, Latin /b, d, g/ are fricatives /β, ð. ɣ/ in medial position between vowels, or between a vowel and a liquid, and remain stops elsewhere, including word-initial position.

Spanish	[riβa]	*riba*	< Latin	*ripam*	'bank'
	[seða]	*seda*	< Latin	*setam*	'silk'
	[fueɣo]	*fuego*	< Latin	*focum*	'hearth'

Obstruent lenitions – voicing of voiceless obstruents

1. Verner's law. In early Germanic, medial voiceless fricatives became voiced when the following vowel was stressed, but not when the preceding vowel was stressed.

E.g., Gothic	*fad*[ð]*ar*	< *fa[θ]ár	'father'	(= Sanskrit *pitár*)
but	*bro*[θ]*ar*	< *brá[θ]ar	'brother'	(= Sanskrit *bhratár*)

Note: Verner emphasizes that the syllabic affiliation of the stop is with the preceding syllable, as attested by alliteration practices in Old Norse verse.

2. The Armenian reflexes of Proto-IndoEuropean (PIE) *nt are /nd/ at the end of accented syllables, but /n/ in unaccented syllables, with /t/ deleted later (Olsen, 1989).

n	< *nt	*-Vn* 3rd p'	< *-nti e.g.	*beren* < *bhéronti 'they bear'
		k'san '20'	< *uikmti	cf. Latin *viginti*
nd	< *nt	*ənderk* 'éntrails'	< *énterah₂	cf. Greek *héntera* 'entrails', Sanskrit *ántara* 'inner'
		and 'field'	< *h₂ánt-o	cf. Sanskrit *ántah* 'boundary' (the semantic shift from 'borders' to 'thing delimited' is attested in other forms in other IE languages)

The Greek reflexes of the PIE forms where the nasal was syllabic, parallel the more general Armenian development, with *nt becoming /ad/ when accented and remaining voiceless /at/ when unaccented. (/a/ is the regular Greek reflex of PIE *n̥.)

Stop insertions

1. In many dialects of Yiddish, an epenthetic stop breaks up [ls] or [lz] clusters in stressed syllables, but not in unstressed syllables.

e.g.,	/álts/	< *als	'everything'	(= standard German *alles*)
	/háldz/	< *halz	'neck, throat'	(= standard German *Hals*)
but	/ʃéjndlz/	< *Sheyndl* (woman's name) + possessive or genitive suffix.		

Note: In the unstressed syllable, the sonorant consonant is syllabic.

2. In English surnames formed by suffixing *-son,* an epenthetic /p/ can break up a heterosyllabic /ms/ cluster after a stressed vowel, but not after an unstressed vowel.

	Thompson	~ *Thomson*	= *Thom* (< *Thomas*) + *son*
	Simpson	~ *Simson*	= *Sim* (< *Simon*) + *son*
	Stimpson	~ *Stimson*	= *Stim* (< *stevm < *Stephen*) + *son*
but	(**Abrampson*)	*Abramson*	= *Abram* + *son*
	(**Williampson*)	*Williamson*	= *William* + *son*